

$H_{0,\mathbf{q}}^G$	Pristine graphene Hamiltonian matrix
μ	Chemical potential
$c_{\mathbf{q}}^\dagger = \begin{pmatrix} a_{\mathbf{q}}^\dagger & b_{\mathbf{q}}^\dagger \end{pmatrix}$	Vector of the creation operators for the carbon p_z orbitals for the two sublattices in momentum space
Δ_{jk}	$D \times D$ matrix coupling the position vectors of i th and j th atoms
Λ_{jk}	$D \times D$ matrix coupling the momentum vectors of i th and j th atoms
$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} & \cdots \\ \Delta_{21} & \Delta_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$	$ID \times ID$ matrix of couplings between the position vectors
$\Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \cdots \\ \Lambda_{21} & \Lambda_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$	$ID \times ID$ matrix of couplings between the momentum vectors
$\epsilon_{s,\mathbf{q}}$	$AD \times 1$ polarization vector from $\mathbf{m}^{-1/2} V_{\mathbf{q}} \mathbf{m}^{-1/2} \epsilon_{s,\mathbf{q}} = \Omega_{s,\mathbf{q}}^2 \epsilon_{s,\mathbf{q}}$

TABLE I. Table of quantities used in calculating the system free energy. Here, A is the number of atoms per unit cell, D is the dimensionality of the system, and I is the number of perturbed lattice atoms

To describe an infinitely-large graphene system with multiple defects, we use the following Hamiltonian:

$$\begin{aligned}
\hat{H} = & \sum_{\mathbf{q}} c_{\mathbf{q}}^\dagger (H_{0,\mathbf{q}}^G - \mu) c_{\mathbf{q}} + \sum_k g_k^\dagger (\varepsilon_k - \mu) g_k \\
& + \sum_{jk} \left[c_{\mathbf{R}_j}^\dagger I_j V_{j,k} g_k + g_k^\dagger (V_{j,k})^* I_j^T c_{\mathbf{R}_j} \right] \\
& + \sum_{jl} c_{\mathbf{R}_j}^\dagger I_j \Delta_{jl} I_l^T c_{\mathbf{R}_l} .
\end{aligned} \tag{1}$$

is a , while $c_{\mathbf{R}}^\dagger$ is its real-space counterpart. g_k^\dagger is the creation operator for the impurity state of energy ε_k . The second line describes the coupling between the impurity states and graphene atoms at unit cells with coordinates \mathbf{R}_j . Importantly, the sum j runs over *all* the atoms impacted by the impurities, either by directly interacting with them or because the

induced lattice deformation changes their coupling to other graphene atoms. To keep track of matrix dimensions, we denote the number of affected atoms by M and the number of impurity states by K . The quantity $I_j^T = \begin{pmatrix} 1 & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & 1 \end{pmatrix}$, depending on the sublattice of the atom j , and $V_{j,k}$ is atom j 's interaction strength with the impurity state k . Finally, the last line gives the perturbation of the graphene Hamiltonian due to the lattice deformation. As with the line above, the sum includes all the modified atoms.

Before proceeding further, we highlight three important aspects of the model. First, what we refer to as the ‘‘impurity state’’ is not just the adatom orbital. Rather, as a consequence of hybridization, it also includes contributions from graphene orbitals not included in the model. As a consequence, the energy ε_k can depend on the carbon-adatom bond length, among other factors, as it influences the magnitude of the orbital interaction. Second, the impurity states are not orthogonal to the graphene Wannier functions due to a finite overlap integral. Equation (1) assumes that the overlap is small and neglects it by treating all the states in the system as orthogonal. While it is possible to extend the treatment to non-orthogonal states, this is outside the scope of our paper. Finally, even though the Hamiltonian in Eq. (1) includes only the carbon p_z orbitals, the subsequent derivation does not depend on this fact. Put differently, to include more orbitals in the model, one simply needs to modify $H_{0,\mathbf{q}}^G$ and adjust the dimensions of I_j in the final result. In this case, the j and l summations run over *orbitals*, not *atoms*.

Using $c_{\mathbf{R}}^\dagger = N^{-1/2} \sum_{\mathbf{q}} c_{\mathbf{q}}^\dagger e^{-i\mathbf{R}\cdot\mathbf{q}}$, where N is the number of unit cells in the system, we write

$$\sum_j c_{\mathbf{R}_j}^\dagger I_j V_{j,k} = \frac{1}{\sqrt{N}} \sum_{\mathbf{q}} c_{\mathbf{q}}^\dagger \underbrace{\left(\sum_j e^{-i\mathbf{R}_j \cdot \mathbf{q}} I_j V_{j,k} \right)}_{\Theta_{\mathbf{q}}^\dagger \mathbf{I} V_k}, \quad (2)$$

where $\Theta_{\mathbf{q}}$ is a column vector of $\mathbf{1}_{2 \times 2} e^{i\mathbf{R}_j \cdot \mathbf{q}}$ for all \mathbf{R}_j , \mathbf{I} is a diagonal matrix of I_j , and V_k is a column vector of $V_{j,k}$. Similarly,

$$\sum_{jk} c_{\mathbf{R}_j}^\dagger I_j \Delta_{jk} I_k^T c_{\mathbf{R}_k} = \frac{1}{N} \sum_{\mathbf{q}\mathbf{q}'} c_{\mathbf{q}}^\dagger \Theta_{\mathbf{q}}^\dagger \mathbf{I} \Delta \mathbf{I}^T \Theta_{\mathbf{q}'} c_{\mathbf{q}'}, \quad (3)$$

where Δ is an $M \times M$ matrix.

Plugging Eqs. (2) and (3) into Eq. (1), we transcribe the Hamiltonian into the imaginary

time action

$$\begin{aligned}
S = & \sum_{\omega_n \mathbf{q} \mathbf{q}'} \bar{\psi}_{\omega_n \mathbf{q}} \overbrace{\left[(-i\omega_n - \mu) \delta_{\mathbf{q} \mathbf{q}'} + H_{\mathbf{q} \mathbf{q}'}^G \right]}^{-G_{i\omega_n + \mu, \mathbf{q} \mathbf{q}'}^{-1}} \psi_{\omega_n \mathbf{q}'} \\
& + \sum_{\omega_n k} \bar{\gamma}_{\omega_n, k} \overbrace{\left(-i\omega_n - \mu + \varepsilon_k \right)}^{-\Gamma_{0, i\omega_n + \mu, k}^{-1}} \gamma_{\omega_n, k} \\
& + \frac{1}{\sqrt{N}} \sum_{\omega_n k \mathbf{q}} \left(\bar{\psi}_{\omega_n, \mathbf{q}} \Theta_{\mathbf{q}}^\dagger \mathbf{I} V_{,k} \gamma_{\omega_n, k} + \bar{\gamma}_{\omega_n, k} V_{,k}^\dagger \mathbf{I}^T \Theta_{\mathbf{q}} \psi_{\omega_n, \mathbf{q}} \right). \tag{4}
\end{aligned}$$

Note that we have combined the \mathbf{q} -diagonal and non-diagonal portions of the graphene Hamiltonian into $H_{\mathbf{q} \mathbf{q}'}^G$. The quantity ω_n is the fermionic Matsubara frequency, and γ and ψ are Grassmann fields. Integrating e^{-S} over all the fields yields the partition function

$$\begin{aligned}
\mathcal{Z} = & \prod_{\omega_n} \left| -\beta G_{i\omega_n + \mu}^{-1} \right| \left| -\beta \left(\Gamma_{0, i\omega_n + \mu}^{-1} - \frac{V^\dagger \mathbf{I}^T \Theta G_{i\omega_n + \mu} \Theta^\dagger \mathbf{I} V}{N} \right) \right| \\
= & \prod_{\omega_n} \left| -\beta \Gamma_{0, i\omega_n + \mu}^{-1} \right| \left| -\beta \left(G_{i\omega_n + \mu}^{-1} - \frac{\Theta^\dagger \mathbf{I} V \Gamma_{0, i\omega_n + \mu} V^\dagger \mathbf{I}^T \Theta}{N} \right) \right|, \tag{5}
\end{aligned}$$

where Θ as a row vector of $\Theta_{\mathbf{q}}$ and V is an $M \times K$ -dimensional matrix. Defining a pristine graphene Green's function $G_z^0 = (z - H_0^G)^{-1}$, we get

$$\begin{aligned}
G_z = & \left[(G_z^0)^{-1} - \frac{1}{N} \Theta^\dagger \mathbf{I} \Delta \mathbf{I}^T \Theta \right]^{-1} \\
= & G_z^0 + \frac{1}{N} G_z^0 \Theta^\dagger \mathbf{I} \Delta (1 - \mathbf{I}^T \Xi_z \mathbf{I} \Delta)^{-1} \mathbf{I}^T \Theta G_z^0, \tag{6}
\end{aligned}$$

where $\Xi_z = \Theta G_z^0 \Theta^\dagger / N$ with entries $\Xi_z^{jk} = \Xi_z^{\mathbf{R}_j - \mathbf{R}_k}$ and

$$\Xi_z^{\mathbf{R}} = \frac{1}{N} \sum_{\mathbf{q}} G_{z \mathbf{q}}^0 e^{i \mathbf{R} \cdot \mathbf{q}}. \tag{7}$$

G_z is the graphene Green's function including the lattice deformation, but not the effects of the impurity states.

In the parentheses of the first line of Eq. (5), we identify the inverse of the full impurity Green's function, denoted by Γ_z^{-1} :

$$\begin{aligned}
\Gamma_z = & (\Gamma_{0,z}^{-1} - V^\dagger \Lambda_z V)^{-1} \\
= & \Gamma_{0,z} + \Gamma_{0,z} V^\dagger \Lambda_z (1 - V \Gamma_{0,z} V^\dagger \Lambda_z)^{-1} V \Gamma_{0,z}, \tag{8}
\end{aligned}$$

$$\Lambda_z = \mathbf{I}^T \Xi_z \mathbf{I} \left[1 + \Delta (1 - \mathbf{I}^T \Xi_z \mathbf{I} \Delta)^{-1} \mathbf{I}^T \Xi_z \mathbf{I} \right]. \tag{9}$$

Also, from the parentheses of the second line in Eq. (5), we obtain the inverse of the full graphene Green's function, given by

$$\begin{aligned}\mathcal{G}_z &= \left[(G_z^0)^{-1} - \frac{1}{N} \Theta^\dagger \mathbf{I} (\Delta + V \Gamma_{0,z} V^\dagger) \mathbf{I}^T \Theta \right]^{-1} \\ &= G_z^0 + \frac{1}{N} G_z^0 \Theta^\dagger \mathbf{I} D_z \mathbf{I}^T \Theta G_z^0,\end{aligned}\tag{10}$$

$$D_z = \left[(\Delta + V \Gamma_{0,z} V^\dagger)^{-1} - \mathbf{I}^T \Xi_z \mathbf{I} \right]^{-1}.\tag{11}$$

Using Eq. (10), it is possible to calculate the real-space graphene Green's function $\mathcal{G}_{i\omega_n+\mu, \mathbf{R}}^s = N^{-1} \sum_{\mathbf{q}\mathbf{q}'} \langle \bar{\psi}_{\omega_n \mathbf{q}}^s \psi_{\omega_n \mathbf{q}'}^s \rangle e^{i(\mathbf{q}' - \mathbf{q}) \cdot \mathbf{R}}$, where s denotes the sublattice and the correlation functions are the diagonal elements of the $[\mathcal{G}_{i\omega_n+\mu}]_{\mathbf{q}'\mathbf{q}}$ blocks:

$$\begin{aligned}\mathcal{G}_{z, \mathbf{R}} &= \Xi_z^0 + \sum_{jk} \Xi_z^{\mathbf{R}-\mathbf{R}_j} (\mathbf{I} D_z \mathbf{I}^T)_{jk} \Xi_z^{\mathbf{R}_k-\mathbf{R}} \\ &= \Xi_z^0 + \begin{pmatrix} \Xi_z^{\mathbf{R}-\mathbf{R}_1} & \dots \end{pmatrix} \mathbf{I} D_z \mathbf{I}^T \begin{pmatrix} \Xi_z^{\mathbf{R}_1-\mathbf{R}} \\ \vdots \end{pmatrix}.\end{aligned}\tag{12}$$

By taking the k th diagonal entry of $-2\text{Im}[\Gamma_{\omega+i0^+}]$ and $-2\text{Im}[\mathcal{G}_{\omega+i0^+, \mathbf{R}}^s]$, we obtain the spectral functions for the k th impurity and the corresponding carbon atom, respectively. These spectral functions can be compared directly to the DFT-computed PDOS. By integrating the Green's functions along the complex axis, we can obtain the occupation number of a graphene orbital or the impurity state. Finally, using the fact that the Helmholtz free energy $F = -T \ln \mathcal{Z}$, one can obtain the impurity interaction energy.